
Recherche des modèles fréquents, corrélations et associations

Exercice 1 Soit la table des transactions suivante :

TID	Liste des items	TID	Liste des items
T100	I1, I2, I5	T600	I2, I3
T200	I2, I4	T700	I1, I3
T300	I2, I3	T800	I1, I2, I3, I5
T400	I1, I2, I4	T900	I1, I2, I3
T500	I1, I3		

1. Représenter cette table par une table formelle.
2. Trouver, en utilisant l'algorithme Apriori, les motifs fréquents sachant que $\sigma_s = 2/9$.
3. Calculer à partir du motifs fréquent le plus long, les règles d'associations les plus fortes sachant que le seuil de confiance est de 70%.

Exercice 2 Soit la base de données transactionnelle suivante :

TID	Items	TID	Items
T01	A, B, C, D	T05	B, C, G
T02	A, C, D, F	T06	D, F, G
T03	C, D, E, G, A	T07	A, B, G
T04	A, D, F, B	T08	C, D, F, G

En utilisant un support minimum de 25% et un minimum de confiance de 60%, trouver :

1. Les motifs fréquents maximaux
2. Les motifs fréquents fermés
3. Les règles solides d'association.

Exercice 3 Supposons qu'il existe 6 articles numérotés de 1 à 6 (a_1, \dots, a_6), et 12 paniers numérotés de 1 à 12 (p_1, \dots, p_{12}). L'article a_i est dans le panier p_j si et seulement si i divise j , par conséquent, a_1 est dans tous les paniers et a_2 est dans les paniers de numéro pair, et ainsi de suite. Avec un support minimum de 25% et une confiance minimale de 90%, donner :

1. La base de données formelle.
2. L'ensemble des motifs fréquents.
3. Les motifs fréquents fermés et les motifs fréquents maximaux.
4. Les règles solides.

Exercice 4 Dans un supermarché, on dispose de la base de transactions suivante :

TID	Items	TID	Items
T1	Lait, Jus, Couches	T6	Lait, Couches, Pain, Beurre
T2	Pain, Beurre, Lait	T7	Pain, Beurre, Couches
T3	Lait, Couches, Sucre	T8	Jus, Couches
T4	Pain, Beurre, Sucre	T9	Lait, Couches, Pain, Beurre
T5	Jus, Sucre, Couches	T10	Jus, Sucre

En utilisant l'algorithme Apriori avec un support minimum de 20% et une confiance minimale de 75%, trouver :

1. Les motifs fréquents,
2. Les motifs fréquents fermés,
3. Les motifs fréquents maximaux,
4. Les règles solides d'association de type $A, B \Rightarrow C$.

Exercice 5 Une banque dispose des informations suivantes sur un ensemble de clients :

Client	M	A	R	E	I
01	moyen	moyen	village	oui	oui
02	élevé	moyen	bourg	non	non
03	faible	âgé	bourg	non	non
04	faible	moyen	bourg	oui	oui
05	moyen	jeune	ville	oui	oui
06	élevé	âgé	ville	oui	non
07	moyen	âgé	ville	oui	non
08	faible	moyen	village	non	non

L'attribut client indique le numéro du client ; l'attribut M indique la moyenne des crédits sur le compte du client ; l'attribut A donne la tranche d'âge ; l'attribut R décrit la localité du client ; l'attribut E possède la valeur oui si le client possède un niveau d'études supérieur au bac ; l'attribut I (la classe) indique si le client effectue ses opérations de gestion de compte via Internet.

1. Donner la base de données formelle correspondant à cette base
2. Calculer les motifs fréquents correspondant à un minimum de support $\sigma_s = 0.4$
3. En déduire les motifs fréquents fermés et maximaux
4. Calculer les règles solides correspondant à un seuil de confiance = 0.9
5. En déduire un modèle de décision.

Responsable de la matière :
Dr A.Djeffal
